



# **Community for Data Integration FY18 RFP Lightning Presentations**

February 6, 2018 Ask questions on slido.com, #X615





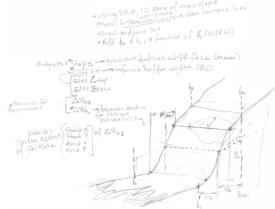
# (A) 4D automated coastal slope failure characterization tool in support of coastal landslide hazard assessments

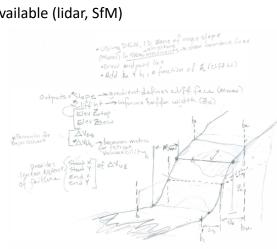
Hapke, Cheryl

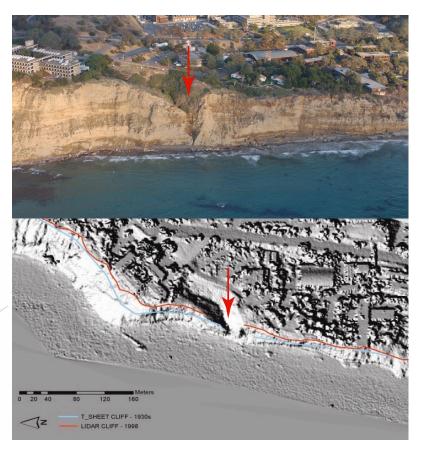
#### 4D Automated Coastal Slope Failure Characterization Tool in Support of **Coastal Landslide Hazard Assessments**

Cheryl Hapke (chapke@usgs.gov)

- Seacliff retreat and coastal landslides are significant hazards along higher relief coasts (CA, PacNW, AK, Great Lakes)
- We are beginning to develop a coastal landslide vulnerability forecast model
- Requirements include metrics of seacliff retreat locations and amounts, and geomorphology
- Traditional methods of seacliff retreat measurement use 2D metrics (i.e. cliff edge positions), primarily manually derived
- 3D datasets are now widely available (lidar, SfM)
- Need is development of an automated approach that utilizes the full value of 3D data and provides more information on cliff change processes over time













# (B) Mapping land-use, hazard vulnerability and habitat suitability using deep neural networks

Warrick, Jonathan



#### **CDI 2018 Statement of Interest**

### Mapping land-use, hazard vulnerability and habitat suitability using deep neural networks (DNNs)



ਤਿ Natural Sciences







#### Jonathan A. Warrick

USGS Pacific Coastal and Marine Science Center

Santa Cruz, CA

jwarrick@usgs.gov

- Oblique or aerial photographs, satellite imagery, orthomosaics
- Non-photographic imagery (e.g. sonar, radar)
- SfM/lidar Workflows
- 3D/4D Geospatial Analyses
- New and legacy data

#### **Automated Semantic**

#### Maps of:

- --Landforms/Sediment/Geology
- --Water (calm/rough)
- --Ground cover/vegetation
- --Specific physical habitats
- --You name it...

#### **Objective:**

Develop classification/mapping workflow for multiple uses:

- --Habitat suitability / biotopes
- --Habitat fragmentation
- --Hazard Vulnerability
- --Resource management



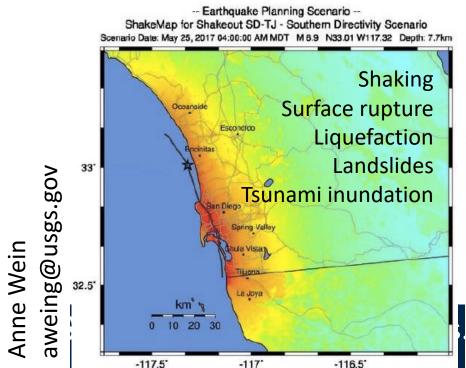




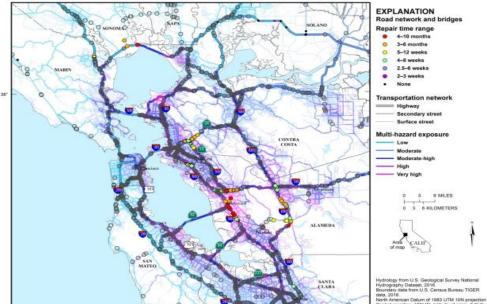
# (C) Multiple hazard and vulnerability data integration and visualization for investigation of cross-border social and economic issues

Wein, Anne

Multiple hazard & vulnerability data integration & visualization for investigation of cross-border social & economic issues







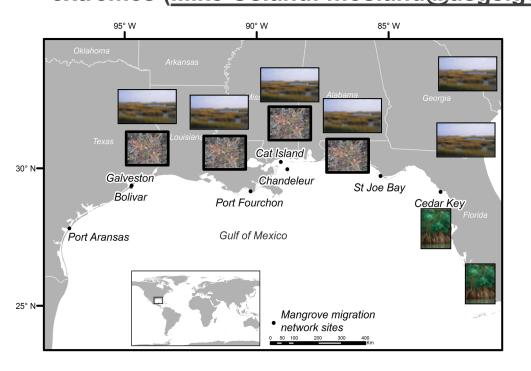




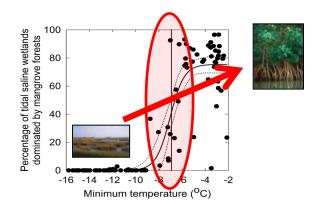
# (D) Mangroves vs. salt marshes: integrating data from a community of practice to assess vulnerability of coastal wetlands to winter temperature extremes

Osland, Michael

Title: Mangroves vs. salt marshes: integrating data from a community of practice to assess vulnerability of coastal wetlands to winter temperature extremes (Mike Osland: mosland@usgs.gov)















# (E) Knowledge Extraction Algorithms (KEA): Turning Literature Into Data

Neilson, Matthew

### Knowledge Extraction Algorithm (KEA):



**Turning Literature Into Data** Literature <u>GeoDeepDive</u> 5,776,536 documents Train / Validate Models **KEA: Machine Learning** and Text Mining python\* **Actionable Database** 

Doc_ID	Species	Waterbody	Latitude	Longitude	MicroCitation
1023-wf24 5-3234	Dreissena polymorpha	Crystal Lake	43.2621	-84.9318	Zebra mussels were found in Crystal Lake, Michigan
1234-ldf24- 3452	Dreossema polymorpha	Lake Pueblo	38.2664	-104.7446	Lake Pueblo tested positive for zebra mussel









# (F) A regional, interactive web-based application to integrate ecological and risk assessment data

Letcher, Benjamin

### A regional, interactive web-based application to integrate ecological and risk assessment data, Ben Letcher, bletcher@usgs.gov

Interactive visualization tools for complex spatial data

#### Culvert risk of failure

Emergency services, transportation disruption, ecological disruption Existing regional database

#### Propose to combine applications

Regional risk of failure Scalable, generalizable Regional (ME to VA) Zoom, Aggregate, Filter Polygons only



ice.ecosheds.org

sce.ecosheds.org

Local (watershed in MA) Zoom, Aggregate, Filter Polygons, Lines, Points











# (G) Workflows to support integrated predictive science capacity: Forecasting invasive species for natural resource planning and risk assessment

Weltzin, Jake

# Workflows to support integrated predictive science capacity: Forecasting invasive species for natural resource planning and risk assessment

#### Goals:

- Produce real-time and short-term forecasts of invasive & pest plant and animal activity
- Deliver products to natural resource managers
- Develop best practices for communication
- Document and share workflow and best practices for reproducibility

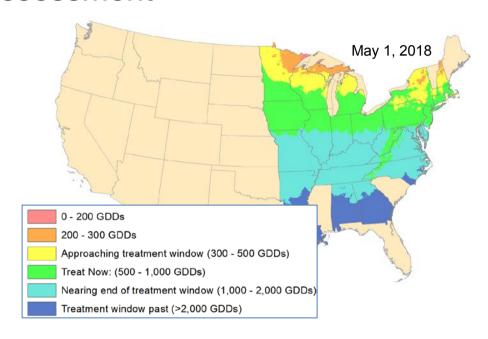
#### Target species for 2018:

- emerald ash borer
- apple maggot

winter moth

hemlock wooly adelgid

- lilac/ash borer
- Canada thistle



Jake Weltzin, jweltzin@usgs.gov





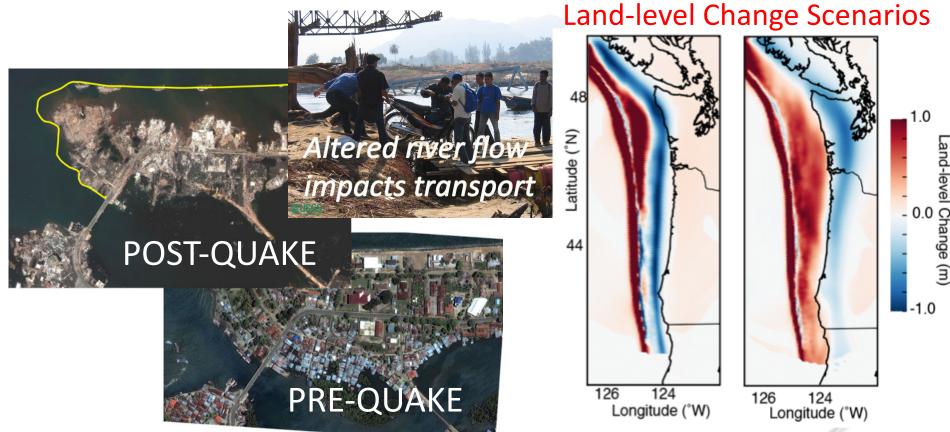




# (H) Assessing and Disseminating Cascading Hazards Information

Gomberg, Joan

#### **Assessing and Disseminating Cascading Hazards Information**











# (I) Delivery of Landsat waterfowl food resource time series maps for drought management in Central Valley wildlife refuges

Byrd, Kristin

### Delivery of Landsat migratory waterfowl food resource time series maps for drought management in Central Valley wildlife refuges

Kristin Byrd, Ph.D. (kbyrd@usgs.gov) and Cynthia Wallace, Ph.D. (cwallace@usgs.gov), Western Geographic Science Center, Menlo Park, CA



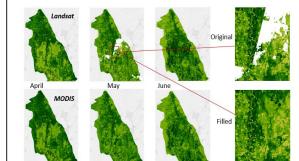
Waterfowl – ducks, geese and swans



Waterfowl food resources produced in Central Valley managed wetlands



Field Data
>250 field plots from 4 USFWS
NWR and 3 State Wildlife Areas



$$M_L = \frac{A_L + J_L}{A_M + I_M} * M_M$$

Prototype method for Landsat greenness anomaly metrics for 2001-2017 using MODIS gap filling



Algorithms in R, Google Earth Engine



Landsat-based time series maps of waterfowl food resource coverage and relative productivity, delivered to managers.









## (J) Content specifications to enable USGS transition to ISO metadata standard

Walworth, Dennis

## Content Specifications to Enable USGS Transition to the ISO Metadata Standard

- ISO is the new standard in town
- ISO: Flexible, but no constraints
- Define USGS Content Specifications
- Easier for metadata authors
- Robust, complete and relevant metadata
- Convene workshop to define USGS Specs

Dennis Walworth (dwalworth@usgs.gov), Fran Lightsom (flightsom@usgs.gov)





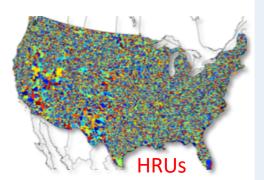




# (K) Developing an integrated predictive forecasting capacity for risk and hazard assessment from local to national scales utilizing the USGS National Hydrologic Model

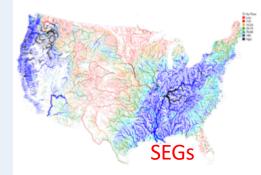
Hay, Lauren

# Developing an integrated predictive forecasting capacity for risk and hazard assessment from local to national scales utilizing the USGS National Hydrologic Model (NHM)



### Ensemble Forecasting (EF) techniques applied to the NHM

- 109,951 Hydrologic Response Units (HRUs)
- 56,460 stream segments (SEGs)



NHM-PRMS run in EF mode for CONUS → produce 90 day forecasts on the first day of every month for the historic record for water budget components and stream temperature forecasts.



#### Forecasts will be analyzed for

- accuracy
- conditioning by atmospheric and sea-surface teleconnections
- ability to predict historic identified risks and hazard events











## (L) Herbarium Specimen Inventory and Archive

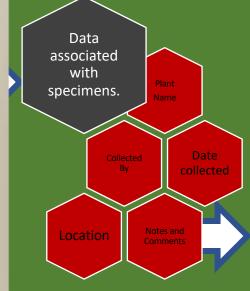
McCoy, John

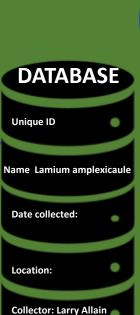
### **Herbarium Specimen Inventory and Archive**

John W. McCoy and Larry Allain

mccoyj@usgs.gov

Collection





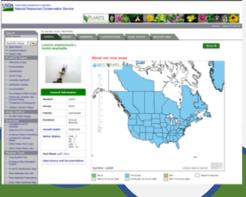
ITIS Data Access

ITIS Advanced Search and Report

Filter on: records with record credibility rating of all ratings



**Associated Databases** 



for collection, photographs and observation.

Data archives, Researcher Resources, Collaboration, Current information







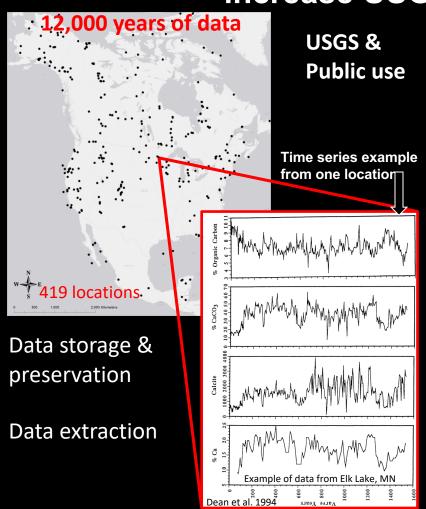


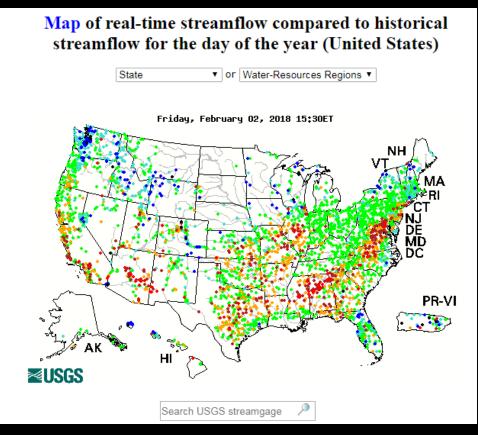
# (M) Integrating and visualizing multiproxy time series data to enhance data organization and accessibility for USGS researchers and increase USGS data access for public use

Rodysill, Jessica



Integrating and visualizing multiproxy time series data to enhance data organization and accessibility for USGS researchers and increase USGS data access for public use





Map output similar to USGS WaterWatch (above) ....but for the past

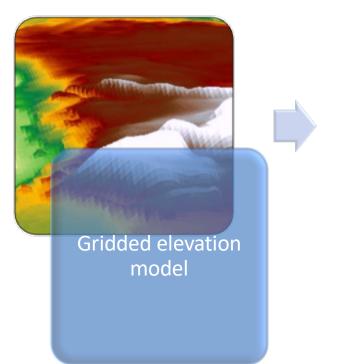


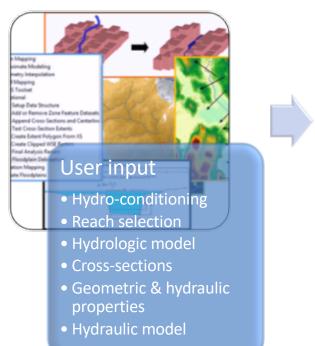


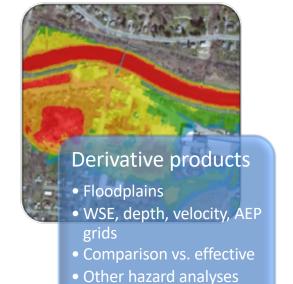
# (N) Floodplain Modeling, Analysis, and Mapping

Schalk, Luther

# Floodplain Modeling, Analysis, and Mapping Luther Schalk, Ischalk@usgs.gov











• Program-specific schema





## (O) ICE! Ice Jam Hazard Mobile-Enabled Website

Chase, Katherine

#### ICE! Ice Jam Hazard Mobile-Enabled Website

- Ice jams MAJOR HAZARD
- Vulnerable sites
- Existing USGS website
- USACE database
- NWS, Montana, Nebraska
- Katherine J. Chase, Hans Vraga, Tonia Roddick

kchase@usgs.gov, hvraga@usgs.gov, troddick@usgs.gov











# (P) Creation of a web-application for distributing unarchived fires

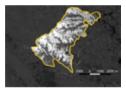
Picotte, Joshua

## Creation of a web-application for distributing unarchived fires

Joshua Picotte ASRC Federal InuTeq LLC, Contractor to the U.S. Geological Survey (USGS), Earth Resources Observation and Science (EROS) Center, Contract Number G13PC00028

#### 1. Bundle







#### 2. Database



#### 3. Serve









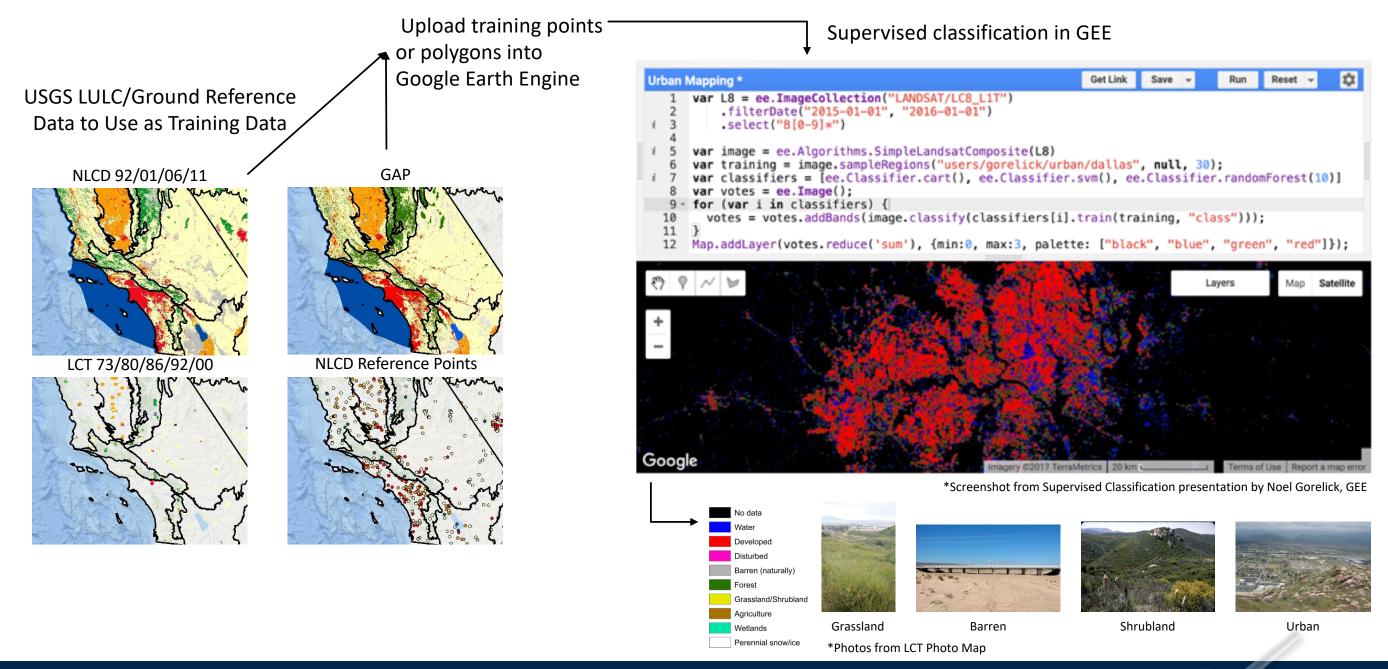


# (Q) Training a Google Earth Engine land cover classification with ground reference data from USGS Projects

Aneece, Itiya

#### Training a Google Earth Engine land cover classification with samples from USGS projects

Itiya Aneece (ianeece@usgs.gov), Jess Walker (jjwalker@usgs.gov), Chris Soulard (csoulard@usgs.gov) Western Geographic Science Center









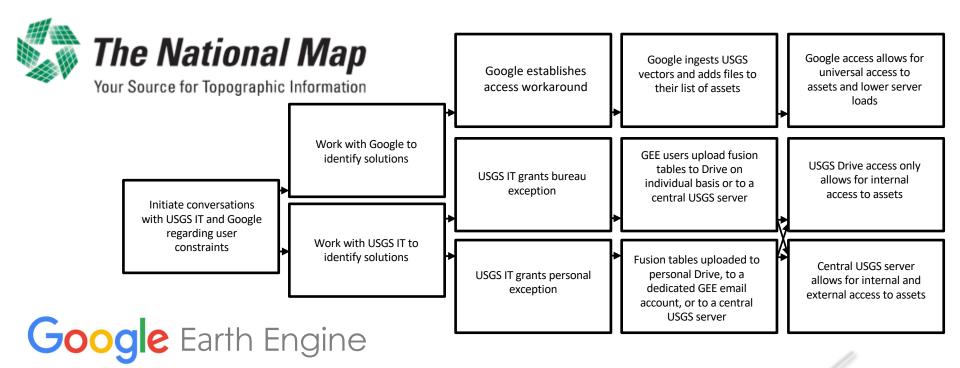


# (R) Facilitating USGS data integration into Google Earth Engine for geospatial analysis

Walker, Jessica

#### Facilitating USGS data integration into Google Earth Engine for geospatial analysis

Jessica Walker (jjwalker@usgs.gov), Itiya Aneece (ianeece@usgs.gov), Christopher Soulard (csoulard@usgs.gov) Western Geographic Science Center











### (S) An Interactive Web-based Tool for Anticipating Long-term Drought Risk

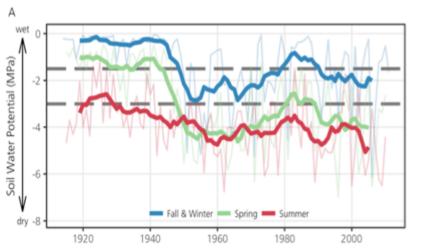
Bradford, John

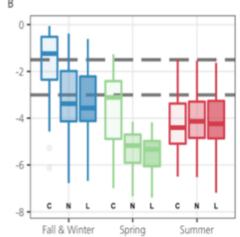
### An Interactive Web-based Tool for Anticipating Long-term Drought Risk

**Need:** Drought risk is expected to increase in coming decades, but no tool exists to predict and understand sitespecific risk of future drought.

#### **Process:**

- Utilize Rshiny and AmazonS3 to create a userfriendly, interactive, online front-end for
- 2. SOILWAT2, a proven, existing model of ecosystem water balance.





#### **Deliverables:**

Site-specific *data*, *metadata* and synthesized quantifications and visualizations of:

- A. long-term historical moisture patterns
- B. predicted **N**ear and **L**ate century moisture





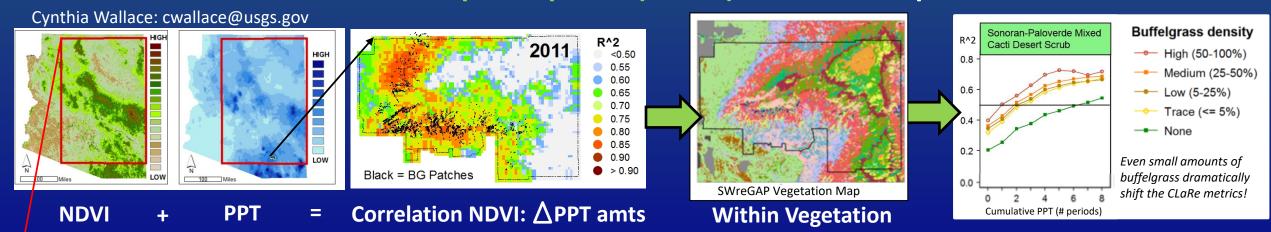




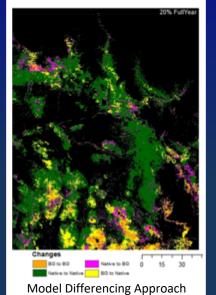
### (T) Tools based on Regional Climate-Landscape Response (CLaRe) metrics for invasive species and fire fuels hazards management

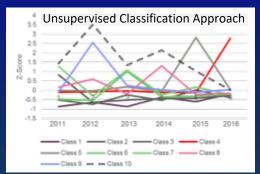
Wallace, Cynthia

### Tools based on Climate-Landscape Response (CLaRe) for invasive species and fuels hazards



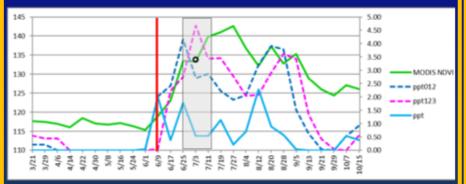
### 1. NASCENT POPULATIONS





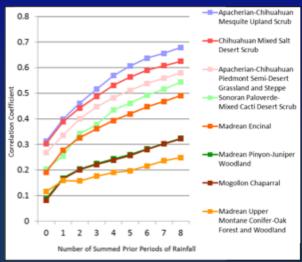
- Map Regionally
- Field validation
- Isolate seasons
- Explore various invasives dynamics

### 2. OPTIMIZE TREATMENT



- Map Regionally, including:
- NDVI, NDVI Slope, Cumulative prior PPT
- Calibrate to management landscape
- Intersection of areas PREDICTED to turn green and OBSERVED to be green

### 3. FOREST FUELS



- Isolate Forest AOI
- Explore/Exploit Rooting Depth Stratification





### (U) Validation of High Level Remote Sensing Products in Support of Essential Climate Variables

Saleh, Raad

### Validation of High Level Remote Sensing Products in Support of Essential Climate Variables Raad Saleh, rsaleh@usgs.gov

**Background:** High Level Remote Sensing (HL-RS) products, such as Surface Reflectance (SR) and Surface Temperature (ST) are generated using current and archived satellite data spanning over serval decades. HL-RS products are building block in the Essential Climate Variables (ECVs) produced by EROS. These are global 30 meter land cover; burned area (surrogate for fire disturbance); dynamic surface water extent (proxy for lake variables); and snow covered area (a component of snow and ice cover).

**ISSUE:** ECVs support monitoring, analysis, and predictive modeling of vast number of applications related to risk assessment and hazard-exposure that affect people, land, and nature. The quality of the HL-RS products, and in turn the ECVs, affect the validity and accuracy of the risk assessment and quantifying the hazard-exposure. The specific issue we are addressing is currently quality assurance of HL-RS does not include validation of the resulting reflectance values based on comparison with the corresponding physical data. A very important HL-RS product can be rendered worthless if no validation protocols are included to ensure accuracy of the derived values, hence affecting the dependent ECVs.

**Solution:** Establish validation protocols to be embedded in production pipeline, and then report the validation outcome into the associated metadata as a quality assurance. The goal is to ensure the measured SR is the best possible representative of the physical feature on the surface.

**Technical Approach:** Comparing calculated SR to known, controlled independent measurements, then quantifying the difference, then reporting results as part of associated metadata. We will evaluate vicarious, multi-sensor, as well as comparing to pseudo-invariant sites.

**Team:** EROS/USGS, South Dakota State University, and Committee on Earth Observation Satellites

**Deliverables:** 1) Best practices document, 2) Technical methodology, 3) Educational activities and 4) a USGS Open-File Report.

**Cost:** Requested Funding \$43,249. Matching Funding \$26,336.

**Durations**: 6 Months









# (V) Reproducible and Executable Workflows for Digitization, Discovery, and Use of Physical Samples

Latysh, Natalie

### Latysh & Motz – 2018 CDI SOI: Reproducible and Executable Workflows for Digitization, Discovery, and Use of Physical Samples **SUSGS**



Sample acquisition and management



- Type of camera
- Scanning resolution
- Collection management and policies
- Environmental controls
- Databases





- Metadata
- standards Describing physical samples
- Creating web services
- Managing digital files

Data management, discovery, use

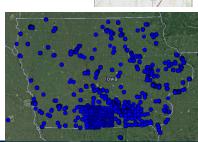


How do I do any of this?

**Jupyter Notebook workflows!** 



KML service showing rock core localities











### (W) Developing a Playbook to Operationalize Crowdsourcing for Emergency Management (EM) through the Standardization of Data, Tools, and Procedures

Liu, Sophia B

### Developing a Playbook to Operationalize Crowdsourcing for Emergency Management through the Standardization of Data, Tools, and Procedures

Sophia B Liu - Innovation Specialist @ Science and Decisions Center (sophialiu@usgs.gov)



#### **Conduct Gap Analysis**

Identify key data points, data standards, tools, and standard operating procedures commonly used for all hazards and disaster phases by the Whole Community.





### **Organize In-Person & Online Forums**

Discuss findings from the gap analysis and develop solutions for standardizing the data, tools, and procedures for crowdsourcing in emergency management.



Humanitarian OpenStreetMap Team







### Develop a Playbook

Incorporate outputs from forums & related initiatives to produce guidance on standardizing the curation and integration of crowdsourced data, tools, & procedures.





### **Iteratively Test & Improve Playbook**

Test playbook to evaluate the utility and fidelity of the crowdsourcing standards at exercises, meetings, and hackathons to implement during 2018 hurricane season.







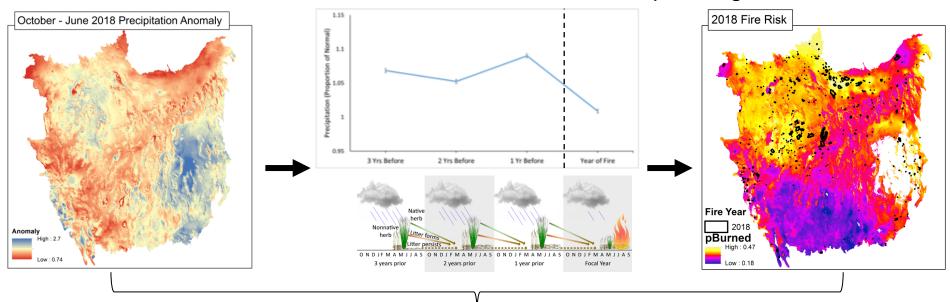


## (X) Creating a wildfire risk prediction tool for public resource managers using antecedent precipitation

Pilliod, David

### Creating a Wildfire Risk Prediction Tool for Public Resource Managers Using Antecedent Precipitation

#### A new tool for fuel treatment and fire risk planning



Plan ahead for an upcoming fire season









### (Y) SfM / Photogrammetry Processing Pipeline on USGS HPC

Adams, Joe

Fort Irwin Bicycle Lake Army Airfield



Cloud Vendors are not FedRAMP Compliant

CDI SOI - Photogrammetry Processing Pipeline on USGS HPC
Joe Adams – jdadams@usgs.gov

Indiana University – Gary Motz Photogrammetry Processing Network



EROS
Data
Sharing
Portal

Legend

♣ Bicycle Lake Army Airfield



30,000 Images 770GBs data

"Data Buddy" Toolbox
Read SD card, Organize Files,
Download UAS Dataflash
Logs, Generate GPX paths
Geotag Photos

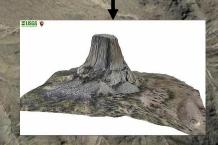
AgiSoft PhotoScan Licenses
No one project can afford to purchase all
needed Licenses. Available to all HPC users.

(4) Front End Processors

USGS HPC Yeti (3) SGI Ultraviolet







Project Study - Stinking Guich Rangely, Colorado

Metadata Offline DB

Phone, Tablet, Laptop
ASC Toolbox?
Like Wunderlist



President Obama Bust

Google Earth

<sub>SanDisk</sub> <u>Ext</u>reme PRO

95 MB/s<sup>-</sup> 55 V30

SanDisk

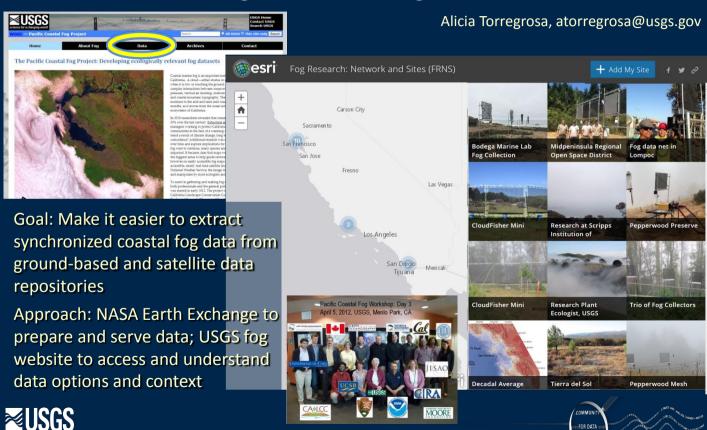




### (Z) Weather and Data Together: A Portal to Coastal Fog Knowledge Generation

Torregrosa, Alicia

### Weather and Data Together: A Portal for Coastal Fog Knowledge Generation



FOR DATA





### (AA) Use of Automated Image Segmentation Algorithm for Improving Classification Workflows of High Resolution Imagery

Oliphant, Adam

### Use of Automated Image Segmentation Algorithm for Improving Classification Workflows of High Resolution Imagery

Adam Oliphant<sup>1\*</sup>, James Tilton<sup>2\*\*</sup>, Thomas Gushue<sup>3</sup>, Nancy Hornewer<sup>4</sup>

- <sup>1</sup> Western Geographic Science Center, <sup>2</sup> NASA Goddard Space Flight Center,
- <sup>3</sup> Grand Canyon Monitoring Research Center, <sup>4</sup> Arizona Water Science Center, <sup>\*</sup> PI, <sup>\*\*</sup> Co-PI

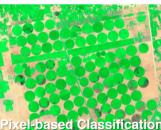
Goal: make tool available to enable researchers to make better use of VHRI

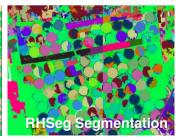
- RHSeg (Recursive Hierarchical Image Segmentation)
- Evaluate non-commercial object segmentation tool on aerial imagery
- Test ability to segment sandbars and vegetation in Grand Canyon

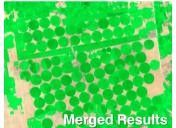
#### **Outcomes**

- Install RHSeg on USGS Nebula Cluster for researches open use
- Present on RHSeg at ESIP Summer Meeting
- Document methods and release products on GCMRC Geospatial Web Server











Papers Utilizing RHSeg doi: 10.3390/rs9101065 10.1109/TGRS.2012.2190079









### (AB) Publicly-Served Wildlife Disease Pathology Atlas Project

Knowles, Susan



### Pathology Atlas

- Collaboration between National Wildlife Health Center (NWHC) and Web Informatics and Mapping (WIM)
- Primarily serves as a means for the NWHC to offer their repository of pathogen information and photos back to the public
- Built with partner submission in mind
- Built using the MEAN stack and hosted in AWS
- Benefits:
  - Supports the core mission of the bureau
  - CDI-specific development artificats

Lead PI: Susan Knowles - sknowles@usgs.gov Presenting: Hans Vraga (Co-PI)





# (AC) Integrating coastal hazards data and models on the semantic web: A proof-of-concept integrated modeling approach

Bagstad, Kenneth

### Integrating coastal hazards data on the semantic web Ken Bagstad, kjbagstad@usgs.gov



#### How to integrate:

- Fast-arriving,
- High-volume
- Unstructured data

into our models?



### Components needed & ready

- Software & server architecture to enable data & model integration
- Semantics: consistent, interdisciplinary terminology to match data & models, reuse existing knowledge wherever appropriate
- Open-source, collaborative, cloud-based, context-aware modeling system using industry standards (W3C)









## (AD) Ontology versioning for unambiguous interpretation of concepts in geoscience data

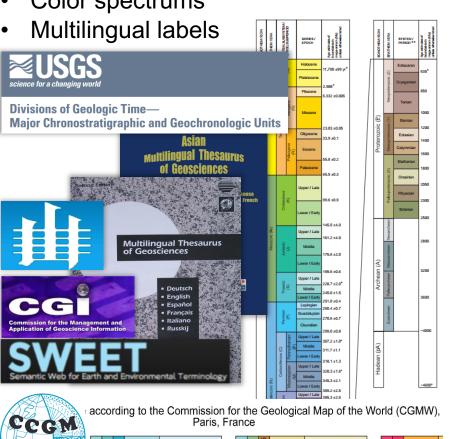
Varanka, Dalia

### **Ontology versioning for unambiguous** interpretation of concepts in geoscience data

Dalia Varanka, CEGIS, USGS dvaranka@usgs.gov Xiaogang Ma, University of Idaho max@uidaho.edu

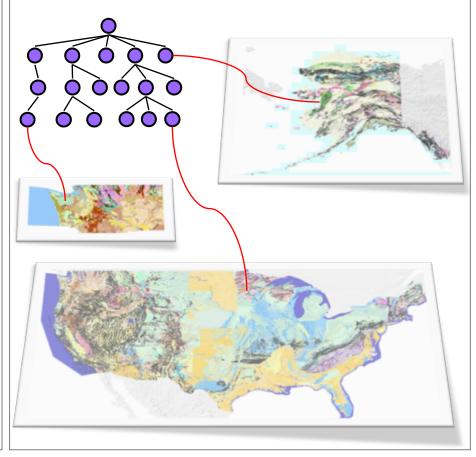
Heterogeneities in geologic time standard and data

- Schema versions
- Color spectrums



### Semantic Web and ontology versioning

- Complete versioning history
- Consistent identifier
- Correct concept meaning







# (AE) Integrating Disparate Spatial Datasets from Local to National Scale for Open-Access Web-Based Visualization and Analysis: A Case Study Compiling U.S. Landslide Inventories

Mirus, Benjamin

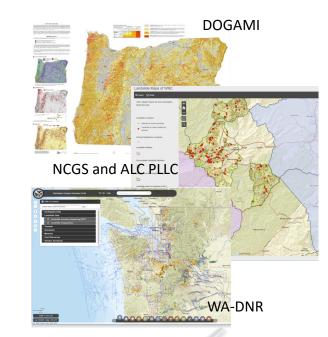
### Integrating Disparate Spatial Datasets from Local to National Scale for Open-Access Web-Based Visualization and Analysis: A Case Study Compiling U.S. Landslide Inventories

National-scale assessment of spatial patterns and trends requires efforts to integrate disparate local-scale data from various agencies

Little guidance or incentive for individual research to collaborate

Proposal leverages existing plans to support a <u>basic</u> repository for landslide inventories across the U.S. to develop:

- 1) Robust visualization and analysis tool of landslide occurrence
- 2) General guidance and framework for integrating and managing disparate spatial datasets for sustainable open-access tools











### (AF) National Alert Risk Mapper for Nonindigenous Aquatic Species

Fuller, Pam

### **National Alert Risk Mapper for Nonindigenous Aquatic Species**

### **Pam Fuller** pfuller@usgs.gov



Specimen ID	1418769
Group	Fishes
Genus	Hypophthalmichthys
Species	molitrix
Common Name	Silver Carp
State	MS
County	Tishomingo
Locality	Bay Springs Lake, lower third
Mapping Accuracy	Approximate
HUC8 Name	Upper Tombigbee
HUC8 Number	3160101
A STATE OF THE STA	4 4 5 4 3 4 5



#### **Silver Carp**

Hypophthalmichthys molitrix



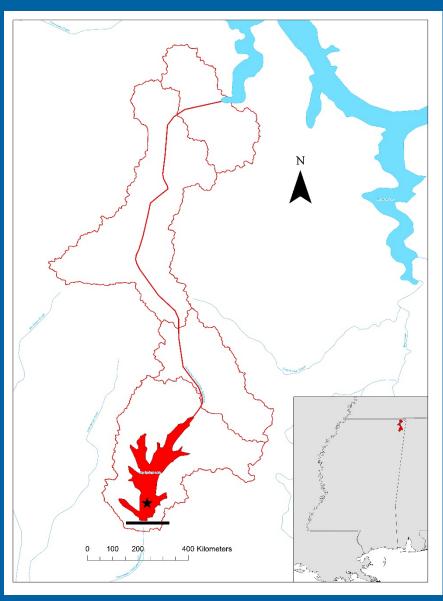


are at risk of



- **Expand ARM to the entire** contiguous US and Hawaii
- Automate the map making process









### How to vote:

See instructions sent by email on Feb. 6 from "GS, CDI"

Email cdi@usgs.gov with any questions.